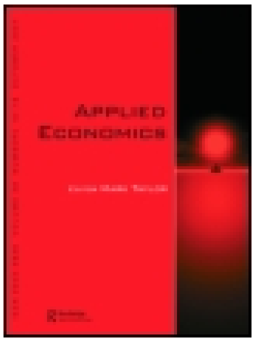

Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: http://www.tandfonline.com/loi/raec20

Geographical disaggregation of sectoral inflation. Econometric modelling of the Euro area and Spanish economies

Gabriel Pino, J. D. Tena & Antoni Espasa

APPLIED ECONOMICS, 2015
http://dx.doi.org/10.1080/00036846.2015.1088141

Geographical disaggregation of sectoral inflation. Econometric modelling of the Euro area and Spanish economies

Gabriel Pino[a], J. D. Tena[b,c] and Antoni Espasa[d]

[a]Universidad de Talca, Facultad de Economía y Negocios, Talca, Chile; [b]Department of Statistics, Universidad Carlos III, Getafe, Spain; [c]Università di Sassari, DiSEA and CRENoS, Sassari, Italy; [d]Departamento de Estadística and Instituto Flores de Lemus, Universidad Carlos III, Getafe, Spain

ABSTRACT
This article studies the performance of different modelling strategies for 969 and 600 monthly price indexes disaggregated by sectors and geographical areas in Spain, regions and in the Euro Area 12 (EA12) countries. We also provide, by means of spatial bi-dimensional vector equilibrium correction models for all pairs of prices between neighbours, a description of spatial cointegration restrictions that could be useful for understanding price setting within an economy. We study the relevance of the regional disaggregation by using the proposed models to forecast the corresponding headline inflation and testing whether it is more accurate than alternative forecasts based on aggregated models. The results for Spain show that this is the case. Country disaggregation forecasts are also reliable for the EA12, but only because derived headline inflation forecasting is not significantly worse than alternative forecasts. The models in this article can be used for competitive analysis and other macro and regional analysis.

KEYWORDS
Spatial cointegration; regional and sectoral prices; regional analysis; relative prices; price setting; competitiveness

JEL CLASSIFICATION
C32; C53; E27; E31; E37

I. Introduction

Recent studies have paid considerable attention to modelling and forecasting a headline rate of inflation by considering information about sectoral disaggregates; see, among others, Hendry and Hubrich (2011). Espasa and Mayo-Burgos (2013) argue that in a variable such as inflation, the aggregate and all its disaggregates matter for policy and investment decisions, and they focus on modelling and forecasting both aggregate and disaggregates, taking some of the common features of the latter into account. There is a large amount of information about the consumer price index of any developed economy, as Statistical Offices provide breakdowns of the respective CPIs by sector and region. A first attempt to use this double disaggregation can be found in Espasa and Albacete (2007), in which the authors work with a breakdown of the euro area Harmonized Index of Consumer Prices (HICP) in just two sectors and five geographical regions, using block-diagonal vector equilibrium correction models, vector equilibrium correction (VeqCM). They find evidence in favour of double disaggregation for forecasting processes. In a more recent paper, Tena, Espasa, and Pino (2010), working with a limited geographical disaggregation for Spanish inflation, find support for the use of disaggregation by sector and region for forecasting headline inflation. Our article extends previous research by considering double disaggregation in two different contexts of economic integration of geographical areas. Thus, we study inflation in the Euro area and Spanish economies using the most detailed information by sector and country or region, respectively. The results could identify significant aspects to consider in further research aimed at formulating a more complex modelling procedure with a more general method for including restrictions between the large numbers of components present in these types of problems.

The vast amount of information derived from disaggregation could also be very useful to understanding the evolution of inflation in a given country or economic area, particularly through the relative performance of price indexes through sectors in different regions or members' states. Consequently, the recent literature on price setting and inflation persistence has focused on disaggregate data. Thus, two contributions by Beck, Hubrich, and Marcellino

CONTACT J.D. Tena Horrillo juande@uniss.it

(2009, 2011) highlight the importance of considering regional factors and a combination of regional and sectoral factors, respectively, for explaining the heterogeneity of disaggregated inflation rates in the euro area. They find that region-specific idiosyncratic components explain a significant part of price variations.

Considering the variation of price series on the regional and sectoral levels is the motivation for this article. However, unlike the aforementioned authors, we evaluate the relevance of the double disaggregation by forecasting the performance of models for disaggregated price data. This is not only something of interest in itself but is also the first step before proposing other economic analysis with this type of models. Two additional contributions are as follows. First, we consider the possible presence of different types of spatial cointegration between neighbouring prices. The question of cointegration between regional sectoral prices could also be interesting in the study of price setting and should, in any event, be considered when modelling inflation rates. Second, here we use the maximum number of sectors, which, in our case, are 50 and 57 for the euro area and Spain, respectively. The problem with using intermediate aggregates from these basic sectors is that common features present between the sectors could disappear in *a priori* definitions of broader sectoral intermediate aggregates. Thus, working with basic sectors is important for avoiding the problem of aggregation bias that could result from grouping statistically heterogeneous sectors together.

Providing forecasts of disaggregated price indexes is valuable in itself, as it enables central bankers and entrepreneurs to identify how different types of economic shocks affect different sectors along regions, with a view to designing an efficient monetary policy, making investment decisions or receiving valuable signals about a possible lack of competitiveness. These economic agents need to know not only this detailed information from the past, but also what it forecasts in individual and relative terms. In this article, we formulate a forecasting procedure for all the disaggregates of a macro-variable, such as inflation at the highest level of breakdown by sector and geographical area. The procedure itself is important because we could easily have around 1000 disaggregates to forecast.

Modelling a large number of components will be useful if their forecasts are reliable. Although this can be tested for each component, it does not seem sufficient. As the components add up to an aggregate, we must test whether the forecast of the aggregate obtained by aggregating the forecasts of the components is, at least, not significantly worse than the alternative forecasts of the aggregate using the same or smaller information sets. This is what we propose in this article.

The structure of this article is as follows. Section II describes and analyses the main features of the time series used in the article. Section III presents and discusses the different methodologies considered for modelling inflation in Spain and the Euro Area 12 (EA12). A discussion of the forecasting results of those methodologies can be found in Section IV. Some concluding remarks follow in Section V.

II. Data description

We use both aggregate price indexes and information related to different sectors and geographical areas. More specifically, we consider the following series: (1) the aggregate HICP for EA12 and the Spanish Consumer Price Index; (2) price indexes for 50 sectors in the EA12 and 57 sectors in the Spanish economy; (3) aggregate price indexes for each of the EA12 countries and the 17 Spanish regions and (4) disaggregated sectoral price indexes for the EA12 countries (600 series) and for 17 regions in Spain (969).[1] Price series for the different Spanish regions (aggregated and disaggregated by sectors) are available from the Spanish Statistical Office (http://www.ine.es). At the European level, disaggregated price series by sectors and countries were obtained from the European Commission (http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home). Spanish series cover the period from 1993:01–2009:12, while EA12 series are available from 1996:01 to 2009:12. We use data up to 2005 to estimate the models and the remaining four years

[1]A description of the sectors, regions and countries can be found in the Appendix. In fact, in Spain, there are 18 regions since the autonomous cities of Ceuta and Melilla can be considered another region. We have been working with these 18 regions, but for the purpose of simplification, we report only the results for the first 17 regions mentioned in the article.

(2006:01–2009:12) to compare the forecasts obtained under different strategies.

For the EA12, Eurostat offers weights of the different countries in order to map the aggregate inflation rate in the EA12 with the national inflation series. However, for Spanish regions, this information is not available from the Spanish Office for National Statistics (INE). We solve the problem by using as weights each region's share of total Spanish expenditures. Indeed, the inflation series obtained by this aggregation is almost identical to the official Spanish inflation rate. Weights at the sector level, on the other hand, are available from the INE and Eurostat. These institutions formulate the aggregate price index for each region based on a chain Laspeyres price index in both cases. During the forecasting exercise, we aggregate inflation projections by using weights computed with information up to the last available period.[2]

In Tables 1 and 2, we report descriptive statistics similar to those in Beck, Hubrich, and Marcellino (2011) for Spain and the EA12, respectively. In general – and consistent with Beck, Hubrich, and Marcellino (2011), Imbs et al. (2005) and Pesaran and Smith (1995) – disaggregated inflation shows low levels of persistency. This could indicate that persistence of aggregated inflation results from aggregation bias that is generated by aggregating heterogeneous price series. A second fact that we observe for both Spain and the EA12 is that there is more heterogeneity across sectors than across geographical areas. Also, the last column of each table indicates a relatively higher degree of co-movement between regions for a given sector than among different sectors for a given region or country. This enhances the importance of taking into account links by sectors in different geographical areas in order to accurately capture the dynamics of disaggregated series. Moreover, fairly heterogeneous values for the mean and volatility for each of disaggregate series suggest the convenience of disaggregated models by sectors and regions for creating a complete picture of Spanish and European inflation.

Figures of price series by level are not shown to save space; however, their inspection reveals that most of them grow smoothly during the period under consideration. Series in first differences, on the other hand, show regular crossing points and no obvious trend. Additionally, many series – for example, prices of lamb, fish, potatoes, vegetables, package holidays, accommodation services, etc. – exhibit a clear seasonal behaviour.

For a formal test on the number of unit roots in the series, we employ the methodology proposed by Osborn et al. (1988) (OCSB henceforth), who extended the procedure of Hasza and Fuller (1982) to seasonal time series for monthly data. Although we are aware of other, more sophisticated procedures to investigate the presence of seasonal unit roots, such as the tests proposed by Franses (1991) and Beaulieu and Miron (1993), we choose the OCSB test because its simplicity enables us to determine whether or not to take seasonal differences instead of testing unit roots one by one at each of the harmonic frequencies of the seasonal cycle.

Results of the test for the disaggregate prices indicate that at the 5% confidence level, the majority of the price series require only one regular difference (and no seasonal differences) to become stationary. For example, at the 5% significance level, results of the tests indicate that for the five biggest Spanish communities – Andalusia, Catalonia, Madrid, Basque Country and Valencia – 77%, 77%, 75%, 72% and 74% of their sectors, respectively, can be considered integrated of order 1 (the average of this proportion for the 17 Spanish communities is 77%). Moreover, at the same confidence level, in the EA12 countries, these percentages are 88%, 80%, 84% and 80% for Germany, Spain, France and Italy, respectively, representing about the 80% of the weighting in the inflation of the EA12 (the average of this proportion for the 12 countries is 85%). Also, in the OCSB equation, the null of insignificant seasonal dummies is rejected at the 5% level in 47%, 49%, 53%, 53% and 42% of the series in Andalusia, Catalonia, Madrid, Basque Country and Valencia,

[2]In the Spanish case, the aggregate for the Ceuta and Melilla regions has been broken into two since 2007, and, therefore, it is not possible to have the complete series. Hence, given the low weight of these two autonomous cities that represent only 0.2% of the total national expenditure, we restrict our analysis to the aggregated price index for Ceuta and Melilla in all the cases by aggregating both series since 2007, according to the share in the total Spanish expenditure. In the case of the EA12 Consumer Price Index, the only irregularities are for the series of education in Belgium and other major durables for recreation and culture in Austria that are available only from 1999:12. Therefore, in these two cases, models are specified and estimated using the information available from that date. Also, other major durables for recreation and culture in the case of Spain are available only from 2006:12. We drop these from the analysis and rescale weights for this fact.

Table 1. Descriptive statistics: Spanish inflation, disaggregation by sectors and regions.

	Level		Volatility		Persistence			
	Mean	Std	Mean	Std	Mean	Std	Disp	Corr(xi,x)
Overall inflation	2.93	0.15	14.77	1.89	0.20	0.03	17.44	0.20
Autonomous communities								
Andalusia	2.81	1.53	14.43	16.39	0.23	0.18	17.35	0.20
Aragon	2.92	1.57	16.37	15.44	0.18	0.19	19.34	0.20
Asturias	2.90	1.57	16.75	18.05	0.17	0.17	20.13	0.16
Balearic Island	3.04	1.59	14.47	10.61	0.14	0.21	16.01	0.23
Canary Island	2.69	1.57	13.10	11.05	0.19	0.17	14.65	0.25
Cantabria	2.84	1.46	16.05	14.00	0.16	0.19	18.95	0.21
Castilla y Leon	2.87	1.54	14.79	15.23	0.25	0.17	18.18	0.20
Castilla la Mancha	2.88	1.48	15.84	16.74	0.22	0.17	19.10	0.21
Catalunya	3.17	1.45	14.56	14.92	0.20	0.19	17.11	0.19
C. Valenciana	2.87	1.50	14.57	15.38	0.19	0.18	17.25	0.21
Extremadura	2.72	1.60	16.33	16.19	0.15	0.19	19.05	0.20
Galicia	2.92	1.41	15.58	16.11	0.23	0.17	17.82	0.21
Madrid	2.83	1.58	13.49	12.59	0.17	0.19	15.89	0.20
Murcia	3.14	1.49	17.70	18.67	0.15	0.20	20.86	0.21
Navarra	3.15	1.47	17.79	15.50	0.15	0.21	20.24	0.18
Basque Country	3.08	1.57	14.92	15.43	0.18	0.17	17.87	0.18
Rioja	3.22	1.44	19.71	22.30	0.16	0.17	23.02	0.14
Sectors								
s1	2.16	0.32	6.54	1.40	0.07	0.13	5.04	0.64
s2	4.15	0.66	10.49	2.81	0.23	0.15	7.22	0.64
s3	3.29	0.54	12.43	3.18	0.27	0.13	9.08	0.68
s4	3.80	0.67	55.78	9.49	0.46	0.07	23.51	0.89
s5	1.85	0.44	28.97	6.18	0.34	0.06	11.86	0.90
s6	2.40	0.36	49.93	11.45	0.10	0.06	18.42	0.91
s7	1.95	0.37	6.59	1.56	0.18	0.14	5.00	0.67
s8	2.21	0.37	42.32	11.36	−0.08	0.06	18.63	0.88
s9	3.26	0.52	12.75	3.16	0.03	0.09	10.06	0.61
s10	2.76	0.65	20.27	5.62	0.27	0.13	13.46	0.76
s11	2.51	0.28	14.97	1.35	0.53	0.09	5.92	0.90
s12	2.10	0.26	9.34	1.68	0.13	0.10	6.05	0.77
s13	3.13	0.20	30.31	2.30	0.57	0.06	10.33	0.92
s14	4.41	0.49	10.62	1.96	0.69	0.04	3.98	0.93
s15	3.70	0.43	10.49	2.67	0.29	0.17	7.96	0.66
s16	4.63	0.35	13.09	1.72	0.57	0.04	4.93	0.92
s17	2.20	0.32	8.62	1.88	0.12	0.15	6.83	0.62
s18	5.10	0.83	86.71	19.70	0.43	0.07	41.26	0.85
s19	2.83	0.37	17.36	2.65	0.49	0.13	9.37	0.82
s20	0.73	0.53	9.37	2.57	0.16	0.18	6.80	0.66
s21	2.39	0.22	4.90	1.26	0.11	0.09	4.08	0.60
s22	1.53	0.51	10.43	2.37	−0.03	0.12	7.87	0.61
s23	2.82	0.36	7.65	2.84	0.20	0.14	5.54	0.66
s24	6.79	0.13	22.04	1.13	−0.01	0.08	2.85	0.96
s25	1.91	0.42	47.06	6.33	0.28	0.04	7.67	0.98
s26	1.82	0.58	59.67	6.70	0.23	0.03	8.74	0.98
s27	1.81	0.80	69.31	11.60	0.16	0.04	10.87	0.98
s28	2.66	0.35	33.58	7.45	0.20	0.06	9.57	0.94
s29	2.56	0.53	34.35	7.12	0.24	0.03	9.32	0.95
s30	2.73	0.70	48.03	5.86	0.21	0.04	10.17	0.96
s31	2.36	0.67	45.85	10.46	0.22	0.04	12.58	0.95
s32	4.30	0.31	6.68	1.79	0.10	0.08	5.17	0.57
s33	4.50	0.36	3.71	1.10	0.33	0.13	2.74	0.62
s34	2.86	0.41	11.96	2.31	0.15	0.08	4.60	0.90
s35	4.12	0.23	5.52	1.71	0.14	0.08	4.39	0.52
s36	2.95	0.45	6.72	2.72	0.17	0.06	3.80	0.76
s37	2.22	0.40	16.78	3.72	0.12	0.05	5.82	0.93
s38	0.21	0.31	3.63	0.92	0.10	0.07	2.86	0.54
s39	2.87	0.26	5.38	1.22	0.11	0.06	4.16	0.60
s40	1.54	0.27	6.70	1.50	0.09	0.10	5.62	0.55
s41	4.36	0.38	6.23	1.21	0.13	0.08	4.30	0.65
s42	3.96	0.39	6.48	1.49	0.12	0.06	3.12	0.82
s43	0.18	0.23	9.42	0.68	−0.04	0.05	2.83	0.95
s44	3.01	0.09	10.71	0.64	0.43	0.01	1.45	0.99
s45	4.83	0.45	12.36	2.91	0.03	0.12	6.91	0.62
s46	4.41	0.46	9.76	2.85	0.20	0.07	4.23	0.87
s47	0.37	0.14	14.08	0.64	0.01	0.00	0.35	1.00
s48	−2.31	0.59	5.69	0.94	0.20	0.13	3.91	0.72
s49	2.75	0.18	5.37	0.41	0.18	0.05	1.71	0.88
s50	2.89	0.48	10.93	4.79	−0.18	0.15	9.60	0.39

(*Continued*)

Table 1. (Continued).

	Level		Volatility		Persistence			
	Mean	Std	Mean	Std	Mean	Std	Disp	Corr(xi,x)
s51	4.71	0.98	12.76	1.95	0.03	0.09	6.12	0.64
s52	4.55	0.64	13.69	4.70	0.06	0.08	4.42	0.81
s53	4.80	0.02	17.48	0.25	−0.07	0.00	0.13	1.00
s54	3.12	0.41	4.73	1.39	0.07	0.10	3.76	0.53
s55	2.85	0.32	4.24	0.76	0.16	0.09	3.31	0.62
s56	4.13	0.21	10.24	2.09	0.11	0.05	3.43	0.95
s57	3.55	0.25	5.36	0.78	0.10	0.07	2.90	0.78

Notes: This table reports descriptive statistics for monthly inflation rates disaggregated by sectors and regions. The reported statistics include the weighted mean and the SD (std) of the time-series means of all inflation series included in a given group (level); the weighted mean and the SD (std) of the time-series SD of all inflation series included in a given group (volatility); the weighted mean and the SD (std) of the persistence measures of all inflation series included in a given group; the average over time of the cross-sectional dispersion of all inflation series included in a given group and the weighted mean of the correlation of all inflation series included in a given group with the group aggregate inflation rate. The measure for persistence is based on the weighted mean of the first-order autocorrelation for all the series.

Table 2. Descriptive statistics: EA12 inflation, disaggregation by sectors and countries.

	Level		Volatility		Persistence			
	Mean	Std	Mean	Std	Mean	Std	Disp	Corr(xi,x)
Overall inflation	1.99	0.48	14.09	9.17	0.12	0.07	20.64	0.15
Countries								
Germany	1.62	1.83	14.47	27.11	0.08	0.23	24.85	0.06
Austria	1.70	1.49	14.19	15.94	0.08	0.19	19.54	0.12
Belgium	1.89	1.52	21.56	38.79	0.11	0.29	34.00	0.14
Spain	2.69	1.81	14.27	19.03	0.29	0.20	20.17	0.16
Finland	1.70	1.80	15.22	15.83	−0.02	0.16	19.84	0.20
France	1.54	1.86	10.88	14.18	0.12	0.25	16.14	0.18
Greece	3.31	1.94	28.26	30.85	0.06	0.21	34.72	0.26
Netherland	1.97	2.18	18.13	23.71	0.05	0.15	26.13	0.17
Ireland	2.45	2.71	15.62	17.81	0.13	0.23	19.13	0.24
Italy	2.26	1.31	11.63	14.97	0.11	0.26	14.19	0.19
Luxembourg	2.32	1.65	14.92	14.28	−0.10	0.24	16.57	0.26
Portugal	2.52	1.65	13.81	15.88	0.14	0.19	18.06	0.13
Sectors								
s1	2.11	0.67	3.76	1.33	0.57	0.16	2.47	0.75
s2	1.87	0.56	5.77	2.78	0.42	0.19	4.82	0.65
s3	2.61	0.47	12.56	5.14	−0.11	0.19	10.43	0.56
s4	1.54	0.61	7.18	2.74	0.55	0.11	3.93	0.72
s5	1.42	0.69	13.55	8.10	0.50	0.19	10.05	0.46
s6	2.26	0.81	37.62	25.83	0.36	0.21	33.17	0.59
s7	1.52	1.42	44.06	22.12	0.29	0.21	32.44	0.75
s8	1.64	0.52	3.98	1.37	0.31	0.22	3.28	0.54
s9	1.61	0.49	3.52	1.84	0.32	0.15	3.45	0.45
s10	0.69	0.53	8.70	5.26	0.44	0.21	7.37	0.67
s11	1.32	0.57	4.68	2.62	0.17	0.11	4.50	0.50
s12	1.41	0.93	5.11	5.89	0.10	0.21	5.17	0.37
s13	1.86	0.78	4.31	2.21	0.18	0.20	3.94	0.51
s14	1.90	0.76	5.50	3.24	0.12	0.18	5.02	0.36
s15	5.23	0.84	17.25	3.00	0.05	0.14	10.33	0.46
s16	0.96	1.24	49.65	33.15	0.06	0.19	36.98	0.83
s17	1.47	1.20	42.81	25.00	0.10	0.17	29.88	0.83
s18	2.54	1.07	3.18	2.65	0.12	0.31	3.42	0.33
s19	2.51	0.77	4.40	1.34	0.10	0.27	3.41	0.48
s20	3.21	0.94	7.26	17.32	0.09	0.16	6.42	0.56
s21	3.15	1.04	15.35	6.07	0.15	0.07	10.33	0.71
s22	1.58	0.73	5.68	4.17	−0.08	0.20	5.40	0.53
s23	1.05	0.99	16.98	14.61	0.02	0.15	16.05	0.62
s24	−0.63	0.74	5.15	8.15	−0.04	0.16	6.64	0.43
s25	2.70	0.50	5.55	2.34	0.06	0.22	4.17	0.40
s26	1.78	0.66	6.67	13.52	0.03	0.13	10.97	0.52
s27	1.23	0.79	3.08	1.94	−0.12	0.18	3.08	0.40
s28	1.05	0.78	3.28	2.16	0.33	0.17	3.42	0.47
s29	2.81	1.16	6.88	3.46	−0.02	0.15	5.44	0.30
s30	2.24	0.68	8.72	3.74	−0.01	0.06	4.55	0.48
s31	0.82	0.62	5.56	2.91	−0.04	0.18	4.58	0.43
s32	0.74	0.56	5.72	3.39	0.03	0.11	5.14	0.42
s33	1.14	0.45	3.87	1.74	0.10	0.12	3.33	0.41

(*Continued*)

Table 2. (Continued).

	Level		Volatility		Persistence			
	Mean	Std	Mean	Std	Mean	Std	Disp	Corr(xi,x)
s34	3.31	0.56	31.51	4.55	0.29	0.14	12.08	0.91
s35	3.29	0.90	4.34	1.81	0.10	0.29	3.65	0.56
s36	2.45	1.15	8.26	7.35	−0.13	0.15	6.54	0.46
s37	2.94	1.06	17.57	7.96	−0.16	0.13	13.22	0.60
s38	2.56	1.66	19.21	19.47	−0.03	0.11	13.81	0.38
s39	−2.60	0.90	10.86	2.29	0.05	0.07	7.59	0.44
s40	−6.13	1.55	7.06	3.98	0.21	0.29	6.41	0.53
s41	1.43	0.71	6.45	4.44	−0.02	0.10	4.51	0.34
s42	0.88	0.52	8.56	4.03	0.09	0.23	7.02	0.64
s43	2.40	0.54	11.03	21.33	−0.09	0.13	7.83	0.36
s44	2.19	0.55	4.40	2.89	0.02	0.10	4.25	0.41
s45	3.58	1.43	86.36	39.08	−0.23	0.26	58.44	0.58
s46	3.14	0.94	12.31	7.09	0.06	0.12	5.42	0.50
s47	2.56	0.84	2.95	3.13	0.18	0.15	3.07	0.58
s48	2.49	1.03	5.44	2.90	−0.01	0.16	4.07	0.50
s49	3.10	0.63	56.82	32.61	0.00	0.25	41.22	0.65
s50	2.23	0.61	3.31	1.40	0.07	0.13	2.81	0.55

Notes: This table reports descriptive statistics for monthly inflation rates disaggregated by sectors and countries. The reported statistics include the weighted mean and the SD (std) of the time-series weighted means of all inflation series included in a given group (level); the weighted mean and the SD (std) of the time-series SD of all inflation series included in a given group (volatility); the weighted mean and the SD (std) of the persistence measures of all inflation series included in a given group; the average over time of the cross-sectional dispersion of all inflation series included in a given group and the weighted mean of the correlation of all inflation series included in a given group with the group aggregate inflation rate. The measure for persistence is based on the weighted mean of the first-order autocorrelation for all the series. The measure for persistence is the first-order autocorrelation.

respectively (the average for the 17 communities is 42%). However, in the EA12 countries, this hypothesis can be rejected in 42%, 58%, 52% and 46% of the cases for Germany, Spain, France and Italy, respectively (the average for the 12 countries is 52%). As a robustness exercise, for the annual rates of inflation in each of the sectors in the Spanish regions and the countries in Europe, we run the Pesaran (2006) panel unit root test, which allows for cross-sectional (spatial) dependence. Results of the test indicate that the null hypothesis is rejected at the 5% level in all the cases for Spain and also for practically all the EA12 series, with only the exception of actual rental for houses. Consistent with this analysis, we specify econometric models in the following sections by assuming that the different price series are generated by unit root processes and by allowing for deterministic seasonality in the cases in which seasonal dummies are jointly significant. However, for robustness, we also consider projections obtained under ARIMA models based on alternative hypotheses about the number of unit roots in the models.

III. Strategies to model regional inflation by sectors in Spain and the Euro area

In this section, we present the strategies to model inflation rates disaggregated by sectors and geographical areas in Spain and the EA12. Then, we evaluate these strategies by their performance in providing an indirect forecast of the corresponding headline inflation during the period 2006:01–2009:12. The forecast evaluation is based on models applied to different degrees of disaggregation. More specifically, for both Spain and the EA12, we compare results obtained from a benchmark strategy, denoted by B, based on a simple ARIMA model specified for the aggregate inflation in Spain and the EA12, with those obtained from a number of alternative strategies that consider different econometric specifications and disaggregation schemes. These strategies can be split into two main groups. The first refers to the use of ARIMA models applied to disaggregated series by sectors and geographical areas in Spain and the EA12. The second approach is based on the specification and estimation of alternative spatial vector equilibrium correction models (SVeqCM), in which the price indexes for each sector can be cointegrated with prices in neighbouring geographical areas. These models use different definitions of neighbourhood based on geographical, economic and sociological considerations, as well as alternative definitions of neighbourhood based on cointegration tests.

The different approaches correspond to different ways to deal with the curse of dimensionality. In fact, under the first strategy, each of the individual time series is restricted to depend only on its own past values, whereas in the second strategy, in addition to past values, we allow for the presence of long-run equilibrium restrictions between prices in the same sector for two neighbouring areas.

In all the cases, we forecast inflation by following a recursive scheme; see, for example, Faust, Rogers, and Wright (2005) and West (2006). Under this approach, the size of the sample used to estimate the parameters of the different models at each forecasting base grows by one observation.

In the remainder of this section, we explain the main features of the two big groups of methodologies used in this article to forecast inflation in Spain and the EA12.

Disaggregated ARIMA models by sectors and geographical areas

The first alternative strategy (A1 henceforth) obtains headline inflation forecasts in the EA12 and Spain from aggregating ARIMA forecasts for each of the 12 European countries and 17 Spanish regions, respectively. Under the second strategy, denoted by A2, we consider sectoral disaggregation only and specify ARIMA models for price indexes in 57 Spanish sectors and 50 sectors in the EA12. The third strategy (A3) considers both sectoral and geographic disaggregation. Thus, we can obtain inflation forecasts in each Spanish region and each European country from the aggregation of forecasts in the different sectors of that specific geographical area, and we can aggregate them again to obtain the headline inflation forecast in Spain and the EA12.

In all cases, our ARIMA models are specified using the TRAMO/SEATS automatic procedure; see Gomez and Maravall (1996).

VeqCM models with spatial cointegration

We also consider VeqCM models to characterize the dynamic pattern for each of the sectoral regional price series. The prototype model takes the form

$$\begin{aligned}\Delta \boldsymbol{p}_{i,j,t} = {} & \boldsymbol{\gamma}_{i,j} + \boldsymbol{\alpha}_{i,j}[\boldsymbol{\beta}_{i,j}{}' \quad \delta_{i,j}]\begin{bmatrix}\boldsymbol{p}_{i,j,t-1}\\ 1\end{bmatrix} \\ & + \boldsymbol{\Phi}_{i,j}\Delta\boldsymbol{p}_{i,j,t-1} + \boldsymbol{\Gamma}_{i,j}D_{1t} + \boldsymbol{\Psi}_{i,j}\mathrm{D}_{2t} \\ & + \boldsymbol{\varepsilon}_{i,j,t} \end{aligned} \tag{1}$$

where $\boldsymbol{p}_{i,j,t}$ is a (2×1) vector containing (logs of) price levels in sector i for a region or country j and its neighbour to be defined; $\boldsymbol{\gamma}_{i,j}$ is a (2×1) vector of intercept parameters; $\boldsymbol{\alpha}_{i,j}$ and $\boldsymbol{\beta}_{i,j}$ are, respectively, the (2×1) adjustment and cointegration vectors; $\delta_{i,j}$ is a scalar that allows for a constant in the cointegration relationship; $\boldsymbol{\Phi}_{i,j}$ is a matrix that accounts for the short-run dynamics; D_{1t} includes centred seasonal dummies, and $\boldsymbol{\Gamma}_{ij}$ is the matrix of parameters associated with these dummies; D_{2t} are centred seasonal dummies that take only non-zero values from 2002:01 to account for the structural break in the seasonal pattern in many disaggregated series in Spain and the EA12 and $\boldsymbol{\Psi}_{i,j}$ is the matrix of parameters associated with this second group of seasonal dummies and $\boldsymbol{\varepsilon}_{ijt}$ is a (2×1) vector of serially uncorrelated errors.

The rationale behind this VeqCM model is very similar to that underlying the Space-Time AR models proposed by Giacomini and Granger (2004). They propose a model that assumes that spatial effects take one period to become manifest and ignores dependence beyond the first temporal and spatial lag. Two important differences between that paper and our approach are that (1) we allow for a cointegration relationship with the neighbour price and (2) we use VeqCM systems with two equations, one for the regional price in question, say P1, and another for the neighbour price, say P2 (instead of imposing neighbouring series as exogenous, as in Giacomini and Granger 2004). These bivariate models are built for all disaggregated price indexes using with their corresponding neighbour prices. In each case, the model is used to forecast P1 only.

We choose the number of lags in Equation 1 to be equal to 1, as this is the specification that minimizes the Schwarz and Akaike criterion in almost all cases in both Spain and the Euro Area. Equation 1 allows for a constant, but not a deterministic, linear trend in the cointegration relationship. This is because a deterministic linear trend in the cointegration relationship amounts to imposing the assumption that prices in the different geographical areas diverge as the forecasting period increases. This specification is not useful for forecasting, as the linear deterministic trend in the cointegration equation can be interpreted as a proxy for other variables not included in the model, and it is reasonable to think that they could be subject to structural shocks during the forecasting period.

We obtain monthly inflation forecasts at the different horizons from Equation 1 by iteration. Then, we compute the annual rate of inflation by adding the 12 monthly rates in the corresponding period.

In many European and Spanish series, there is a structural seasonal break from the period 2002:01 that can be explained by a methodological change in the way that series were collected. Take, for example, the case of different prices for shoes and clothes in both Spain and the EA12. We account for this change in the seasonal pattern by allowing the set of seasonal dummy variables in Equation 1 to have a different impact before and after the break period. Then, using F-statistics, we test each new observation for whether seasonality can be captured with or without a structural break or if there is seasonality at all. In the initial estimation sample, T, and during the forecasting exercise, at each period $T + h$, we run an F-test for deterministic seasonality.

For each sectoral regional price, we build models for the following alternative definitions of neighbours: (1) the whole area (Spain or EA12)(C1), (2) the aggregate of geographical areas with similar economic growth (C2), (3) similar per-capita income (C3), (4) similar macroeconomic conditions (C4) or (5) similar density of population (C5) and (6) the aggregate of geographical neighbours (C6).[3]

We also use two additional definitions of neighbourhood. The first one (C7) is based on the cointegration test proposed by Johansen (1995) and considers that the neighbours for a price index in sector i and region j are the average of all the price indices for that sector in all the other regions for which the null of no cointegration is rejected at the 5% level. The second strategy (C8) defines neighbourhood using ADF tests applied to relative price indices for all the possible pairs of geographical areas in a given sector. Then, we consider as the set of neighbours the average of all the prices for which the null of nonstationarity is rejected at the 5% level.

In the case of the last two strategies, the econometric tests for cointegration and unit roots are repeated for each period during the forecasting exercise. This allows for a flexible definition of neighbours that could be different at different times. In the few cases in which we do not find cointegration under either C7 or C8, we specify an unrestricted bivariate VAR model for variables on first differences.

Note that each of the above definitions imposes a single concept of neighbourhood for all the price indexes across sectors. However, one could assume that different concepts of neighbourhood could be applied to different sectors. In order to account for this fact, we consider another strategy (C9). In it, at each forecasting base, we select the model with the lowest Schwarz criterion between the strategies A3 and C1–C8.

It is also possible that there is some combination of spatial VeqCM and ARIMA models that has not been considered in the previous strategies and could improve the forecast of overall inflation. In order to explore this issue, we define two *ex-post* additional strategies. In the first one (C10), we select for each sectoral regional price index the strategy (A3 and C1–C9) that provides the best individual inflation forecast according to the root mean square forecast error (RMSFE henceforth) and then aggregate all of them to obtain the headline inflation forecast for Spain and EA12. The second strategy (C11) deals with the forecasts of the aggregate sector prices and consists of forecasting inflation in a given sector for Spain or the EA12 using the best strategy. We accomplish this by comparing the RMSFE obtained from the aggregated ARIMA model in strategy A2 and the RMSFE obtained from the best strategy according to all the alternatives A3 and C1–C9. Then, we aggregate the inflation forecasts in the different sectors to estimate the overall rate of inflation in Spain or the EA12. Note that RMSFE under strategies C10 and C11 can be obtained only after inflation data are known. Therefore, they cannot be considered competing strategies, but, rather, as a way to observe the best forecast that could be obtained if the best model were used in each case.

IV. Results

Cointegration analysis, forecasting inflation and relative prices in Spanish regions

One important problem that arises in evaluating the forecast of Spanish inflation is the high degree of

[3]A description of the series contained in the different groups of neighbours for each strategy and all the results not explicitly shown in the article can be obtained from the authors upon request.

Table 3. RMSFE for the Spanish headline inflation of the Benchmark strategy and relative RMSFEs with respect to Benchmark under alternative strategies.

	Period 2006:01–2009:12				Period 2006:01–2008:12			
	1P	4P	8P	12P	1P	4P	8P	12P
B	0.66	1.97	2.93	3.65	0.56	1.51	1.85	2.14
A1	0.91*	0.99	1.03	1.07	0.89*	1.04	1.04	1.10
A2	0.63**	0.71*	0.80	0.84	0.66**	0.77	0.75	0.83
A3	0.62**	0.69*	0.77*	0.82	0.64**	0.75	0.71	0.79
C1	0.60**	0.70*	0.82*	0.87	0.64**	0.77*	0.74*	0.82

Note(s): B: ARIMA model for the aggregate series; A1: ARIMA models applied to regions; A2: ARIMA models applied to sectors; A3: ARIMA models applied to sectors and regions; C1: SVeqCM with the whole area. * denotes rejection at the 0.05 significance level and ** denotes rejection at the 0.01 significance level relative to the benchmark strategy by using the modified Diebold and Mariano (1995) test, as proposed by Harvey, Leybourne, and Newbold (1997).

volatility in the inflation rates after the economic crisis at the end of 2008. Hence, for robustness, we evaluate the performance of the different forecasting strategies for both the periods 2006:01–2009:12 and 2006:01–2008:12. Table 3 shows the RMSFE of the benchmark strategy and the RMSFE of each of the alternative strategies relative to the benchmark. An RMSFE ratio lesser than 1 for a particular strategy indicates an improvement over the benchmark. The table also indicates whether the forecasts are significantly different using the modified Diebold and Mariano (1995) test proposed by Harvey, Leybourne, and Newbold (1997).

1. As expected, the economic crisis had a negative influence on the accuracy of forecasts under all the strategies. However, the main conclusions about the relative efficiency of the different methodologies are unaffected by this consideration. For that reason, the remainder of this section is based on analysis that includes 2009 in all cases. We now discuss our conclusions.

2. We find that geographical considerations (A1) alone are not relevant or even disruptive, while the use of the sectoral disaggregation (A2) on its own implies an improvement in forecasting accuracy, which is significantly different from the baseline forecasts for short horizons.

3. Moreover, compared with the strategies that use only a single disaggregation criterion, strategy (A3), which uses double disaggregation, always improves the headline inflation forecast. In fact, for the whole sample, the modified DM test to compare strategies A2 and A3 takes values of 1.62 and 1.51 for one and four periods ahead, indicating that they are not significantly different at the 5% level but are at the 10% level. For longer horizons, in many cases, the values of these statistics surpass the critical values at the 5% significance level. For example, the values of the statistics for the forecasts 9, 10, 11 and 12 periods ahead are 3.82, 3.61, 3.68 and 2.17, respectively, and the null hypothesis of forecast equality is rejected at the 5% level in all these cases.[4]

4. The results for strategies C1–C9, which are not reported in the table but are available from the authors by request, using spatial cointegration indicates that the different definitions of neighbour used in this article do not lead to significantly different accuracy in determining headline inflation.

5. About whether it is better to use an univariate ARIMA model – strategy (A3) – or a SVeqCM – strategies (C1)–(C9) – to forecast inflation for each of the 969 disaggregated series, we obtain that the best forecasting strategy using cointegration terms is C9. In this strategy, for each sector in each region, we perform a test to determine which definition of neighbour leads to the best SVeqCM according to the Schwarz criterion. In any case, as mentioned in point 4, the performance of C9 is not significantly different from the other Cs' strategies – in particular, C1's – that consider cointegration with the corresponding Spanish sector. Therefore, for simplicity, in what follows, we focus our analysis on the comparison of strategies A3 and C1 with respect to A2 (in the case of sector inflation forecast) and A1 (in the case of geographical inflation forecast). C1 performs better than A3 at horizon 1 but not at the other ones. In both cases, the differences are not significant.

6. Figures 1 and 2 show the best strategy to forecast inflation at horizons 1 and 12, respectively, for each of the individual series and for the sector aggregates (last column). Table 4 summarizes the main results by sectors and regions. For the one-step-ahead forecast, we observe that the best forecast for the aggregate of a region is always obtained

[4]Note that ARIMA models specified with TRAMO/SEATS assume in most cases that price series require a regular and a seasonal difference to become stationary. We also specify univariate models applied to the series with only one regular difference and, when they are significant, seasonal dummies to sectoral disaggregated price series. The results are very similar and the same conclusions maintained. For example, for strategy A2, the RMSFE using ARIMA models with TRAMO are 0.42 and 3.10 at horizons 1 and 12, while the RMSFE using the proposed alternative ARIMA models are 0.42 and 3.08 at horizons 1 and 12. Therefore, for simplicity, only the results with ARIMA models obtained from TRAMO are reported here.

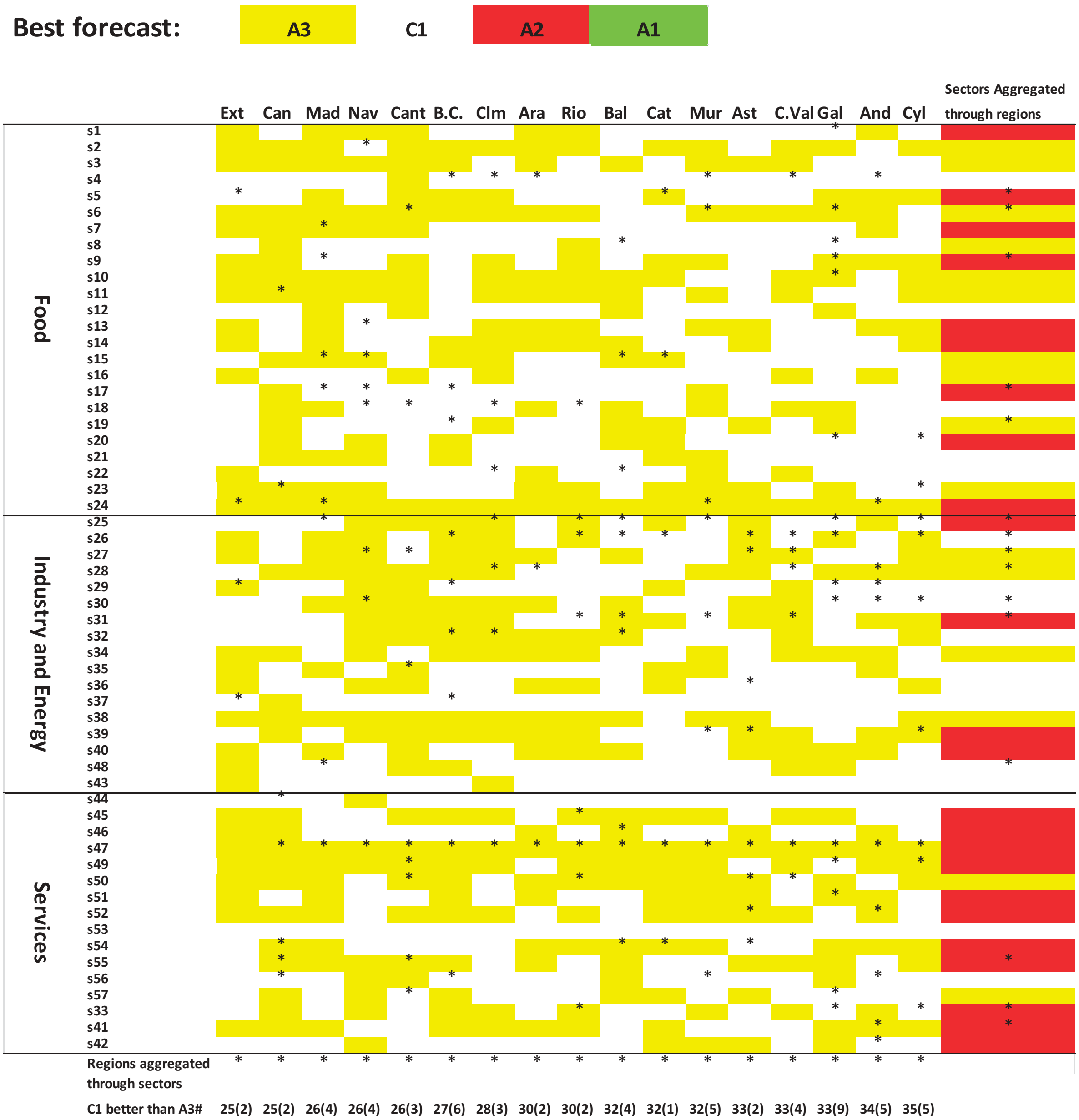

Figure 1. Best forecasting strategy for Spain according to RMSFE. One period ahead.

Notes: * denotes rejection at the 0.05 significance level by using the modified Diebold and Mariano (1995) test, as proposed by Harvey, Leybourne, and Newbold (1997). This test has been run by comparing strategies A3 and C1 for each individual series, and A2 (A1) with respect to the best of A3 and C1 for aggregated regions (aggregated sectors). # Between brackets, the number of cases significant at the 5% level.

under strategy C1. Also, the majority of the 969 inflation series are better forecast by considering cointegration relationships with Spain instead of using ARIMA models. For longer horizons (see Table 4), the opposite is true. This is consistent with Christoffersen and Diebold (1998), who find that VeqCM models are particularly useful to forecast in the short run, as they identify situations of disequilibrium and indicate the dynamic of the model's variables del to return to equilibrium in subsequent periods. In fact, Table 3 shows that for horizon 1, the best procedure to forecast headline inflation is C1.

7. On the other hand, if the purpose of the analysis is to predict inflation in each of the 57 national sectors, the best strategy in more cases (24 of 57) is A2, which uses an ARIMA model for the aggregate national series of each sector. Nevertheless, for forecasting overall inflation (see Table 3), the best

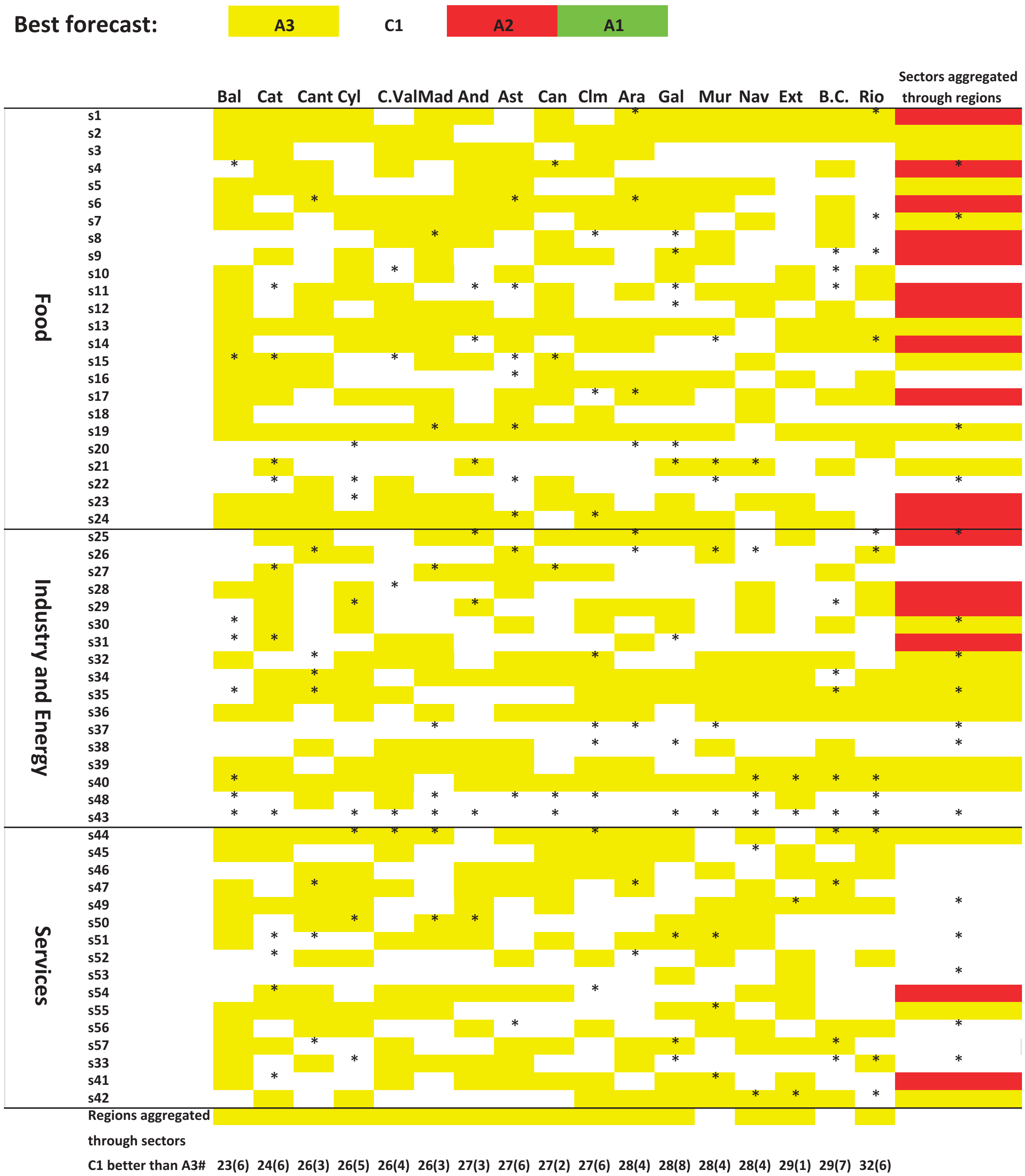

Figure 2. Best forecasting strategy for Spain according to RMSFE. Twelve periods ahead.

Notes: * denotes rejection at the 0.05 significance level by using the modified Diebold and Mariano (1995) test, as proposed by Harvey, Leybourne, and Newbold (1997). This test has been run by comparing strategies A3 and C1 for each individual series, and A2 (A1) with respect to the best of A3 and C1 for aggregated regions (aggregated sectors). ≠ Between brackets, the number of cases significant at the 5% level.

procedure is C1. For longer horizons (see Table 4), the best strategy in more cases is C1, followed closely by A3, as A3 shows better accuracy for the overall aggregate (Table 3).

8. We run a series of robust exercises but do not explicitly report them here for the sake of brevity. First, we also estimate dynamic common factors by principal components that explain most of the

Table 4. Number of cases that the different strategies A2, A3 and C1 show the best forecasting performance in the EA12 and in Spain.

	Spain			Euro Area		
	Sectors	Regions	Disaggregated	Sectors	Countries	Disaggregated
1 Step-ahead forecast						
A2	24(8)			42(32)		
A3	16(4)	0(0)	462(78)	2(1)	2(2)	336(56)
C1	17(3)	17(17)	507(57)	6(2)	10(6)	264(42)
4 Step-ahead forecast						
A2	13(5)			34(20)		
A3	21(5)	9(9)	507(99)	10(1)	3(1)	324(69)
C1	23(0)	8(8)	462(67)	6(0)	9(3)	276(31)
8 Step-ahead forecast						
A2	17(0)			28(9)		
A3	19(1)	15(6)	522(79)	16(2)	3(0)	363(80)
C1	21(2)	2(2)	447(65)	6(0)	8(0)	237(41)
12 Step-ahead forecast						
A2	17(1)			22(4)		
A3	19(2)	15(0)	508(65)	18(3)	3(1)	368(44)
C1	21(4)	2(0)	461(74)	10(1)	5(0)	232(30)

Note: The number of cases in which the strategy is significantly better at the 5% level than the second-best strategy is shown between brackets.

variability of annual inflation series in each of the aggregated sectors. The baseline model is given by

$$\Delta_{12}p_{i,j,t} = c_{i,j} + \sum_{r=1}^{p} \beta_{i,j,r} f_{i,r,t} + e_{i,j,t} \qquad (2)$$

where $c_{i,j}$ is a constant parameter; $f_{i,r,t}$ is a common factor for the ith sector in all the Spanish regions; $\beta_{i,j,r}$ is a factor loading coefficient and $e_{i,j,t}$ is the idiosyncratic component. We estimate models for p = 1, 2, 4 and 8 common factors by principal components for each of the 57 sectors. Then we use the approaches described by Boivin and Ng (2006) and Schumacher and Breitung (2008) to forecast the 969 series, using direct, indirect and unrestricted factor forecasts of sectoral inflation for each model.

The results show that the forecasting strategies based on dynamic factors do not improve the forecast of headline inflation is most cases. The headline inflation forecasts under dynamic factors are significantly worse than those obtained under strategies A3 and A2 at horizons 1 to 4, and they were not significantly better than these forecasts at horizons 5 to 12. For robustness, we also specify dynamic factor models for the monthly rate of inflation, including in these models one or two sets of seasonal dummy variables in the same way that we did for the VeqCM models. Then, we use these models to forecast the annual rate of inflation. The RMSFE under the best forecast with dynamic factors are 0.47, 1.54, 2.39 and 2.96 at horizons 1, 4, 8 and 12, respectively. This does not change the main conclusion of our analysis.

9. The robustifying procedure proposed by Hendry (2006) is applied to all the strategies using bivariate models SVeqCM, but with no important differences with the previous results.

10. In a final set of experiments, for each of the 969 disaggregated prices, we specify single-equation models in which we allow the dependent variable to react to several price differences between the region in question and each of the other regions, in a spirit similar to Aron and Muelbauer (2012). However, this specification does not improve inflation forecasts in most cases.

Cointegration analysis, forecasting inflation and relative prices in the Euro area 12

In the same vein as the Spanish inflation case, Table 5 shows the RMSFE for the benchmark strategy and the relative RMSFE obtained under different strategies. For simplicity, we omit in this table's results from strategies C2–C11, given that they are very similar to those obtained under C1. Table 6 reports a comparative evaluation of strategies A2, A3 and C1, following Harvey, Leybourne, and Newbold (1997). The main conclusions are as follows.

1. Unlike the Spanish case, the strategy based on disaggregation by geographical regions and sectors is not the best one for forecasting EA12 headline inflation, but it does not perform significantly worse than A2. Strategy A2 provides lower,

Table 5. RMSFE of the benchmark strategy and relative RMSFE with respect to benchmark under alternative strategies for the Euro Area 12.

	Period 2006:01–2009:12				Period 2006:01–2008:12			
	1P	4P	8P	12P	1P	4P	8P	12P
B	0.37	1.05	1.73	2.30	0.33	0.83	0.99	1.25
A1	1.12	1.13	1.10	1.08	1.06*	1.07	1.10	1.06
A2	0.77**	0.81*	0.84	0.95	0.70**	0.88	0.89	1.00
A3	0.81**	0.93	0.98	0.99	0.79**	0.93	0.97	0.94
C1	0.78**	0.87*	0.90	0.95	0.73**	0.89	0.90	0.93

Notes: B: ARIMA model for the aggregate series; A1: ARIMA models applied to regions; A2: ARIMA models applied to sectors; A3: ARIMA models applied to sectors and regions; C1: SVeqCM with the whole area. *denotes rejection at the 0.05 significance level and ** denotes rejection at the 0.01 significance level relative to the benchmark strategy by using the modified Diebold and Mariano (1995) test, as proposed by Harvey, Leybourne, and Newbold (1997).

Table 6. Comparison of strategies A2, A3 and C1 for the Euro Area 12.

One period ahead			Twelve periods ahead		
	C1	A3		C1	A3
A2	−0.24	−1.51	A2	0.08	−0.40
C1		−0.84	C1		−0.93

Notes: *denotes rejection at the 0.05 significance level and ** denotes rejection at the 0.01 significance level relative to the benchmark strategy by using the modified Diebold and Mariano (1995) test, as proposed by Harvey, Leybourne, and Newbold (1997).

although not significantly different, RMSFE for all forecast horizons than the alternative disaggregated methodologies, except at horizon 12 in the evaluation period 2006:01 to 2008:12. This implies that the reliability of the disaggregated forecasts for the 600 country sectors is not rejected and that forecasts of the relative sectoral prices between countries could be based on them. This difference in the importance of the geographical dimension in the analysis of Spanish and AE12 inflation data could reflect that spatial links are stronger between the regions of a country than between the countries of an economic union.

2. The procedures in this article allow us to generate inflation forecasts by country, by sector or by both.

When the purpose is to forecast the headline inflation for a particular country, strategies A3 and C1, which break down prices by sector, perform better than the aggregated strategy A1. There are no exceptions in any case for the one-step-ahead forecasts. This superiority of disaggregate models to forecast inflation by country is also evident at longer horizons. For example, for the 12-step-ahead forecasts, strategies A3 and C1 outperform A1 in 8 out of 12 countries.

3. Figure 3 shows the relative performance of strategies A2, A3 and C1 in forecasting inflation in each of the EA12 sectors. At horizon 1, A2 is clearly the best strategy in 44 of 50 sectors in which they are significant at 5% in 32 cases. Inflation forecast could be significantly improved only by considering either A3 or C1 in three sectors.

4. It is also interesting to compare strategies A3 and C1. In forecasting the overall inflation in the EA12, C1 performs better than A3 (see Table 5), but not significantly (see Table 6). In forecasting inflation in each of the 50 sectors of the 12 countries we have that for the one-step-ahead forecast, cointegration improves inflation forecasting compared to simple extrapolative devices in at least 8 out of 12 countries for sectors: s1: bread and cereals; s3: fish and seafood; s5: oils and fat; s14: beer; s38: major appliances and s48: recreational objects. Note that all of them correspond to tradable goods whose prices, due to the possibility of arbitrage, are not expected to diverge through countries.

V. Concluding remarks

In this article, we study the performance of different strategies to model and forecast 969 and 600 monthly price indexes disaggregated by sectors and geographical areas in Spain and the EA12, respectively. We deal with the curse of the dimensionality problem, avoiding modelling with vectors of dimension higher than two. Thus, we specify and estimate ARIMA models, as well as alternative SVeqCM models, where the price index for each sector is allowed to be cointegrated with price indexes in neighbouring geographical areas using different definitions of neighbourhood. The results for both economies show that when disaggregating by just one criterion, sector or region, the former model is more relevant than the latter in forecasting the corresponding headline inflation. These results confirm those of Espasa and Albacete (2007), who use much more reduced disaggregation levels – 10 country sectors in the EA compared with the 600 in this article.

The relevance of the use of the double disaggregation criteria based on sectors and geographical areas seems to depend on the level of economic integration between the areas. Thus, the sectoral breakdown by regions within a country such as Spain improves

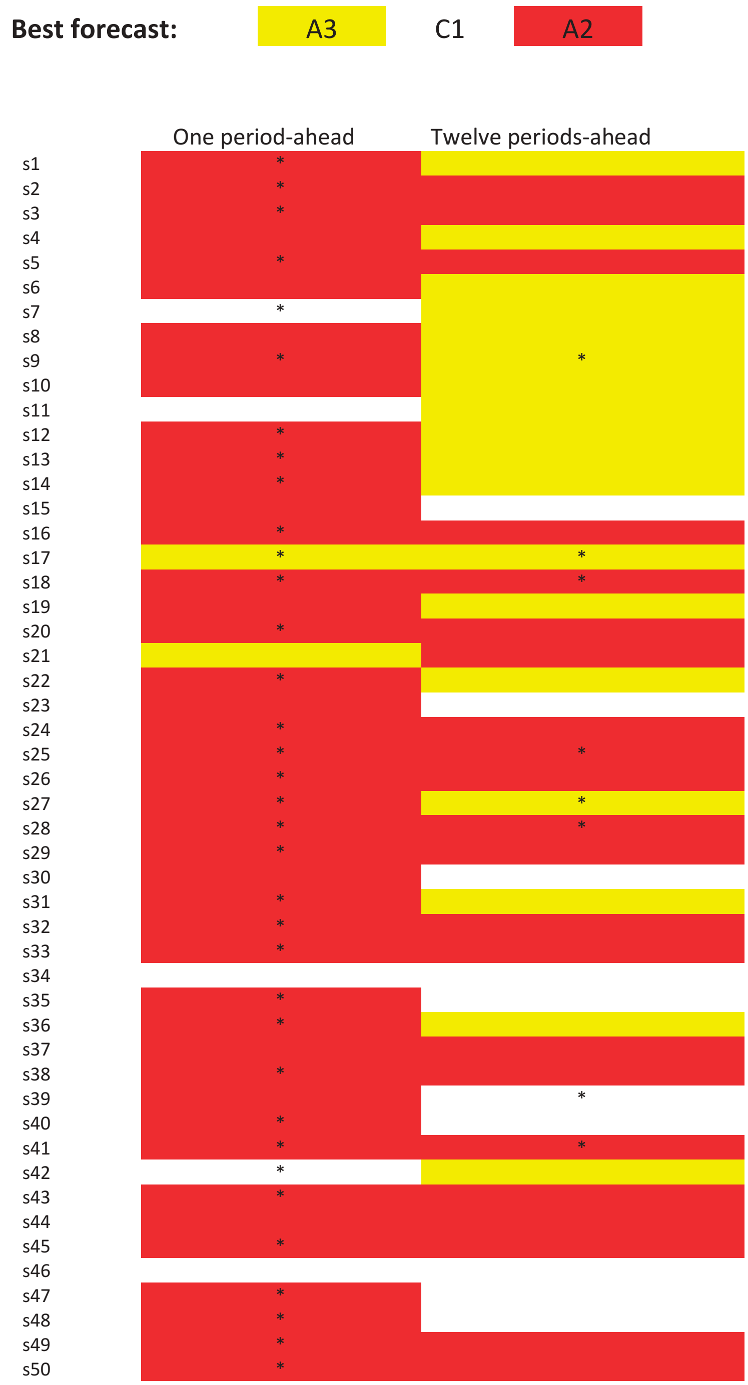

Figure 3. Best forecasting strategy for the Euro Area 12 according to the RMSFE, one and twelve periods ahead.
Note: * denotes rejection at the 0.05 significance level by using the modified Diebold and Mariano (1995) test, as proposed by Harvey, Leybourne, and Newbold (1997).

the accuracy of headline inflation forecasting, but this is not the case when breaking down the sectors by countries in the EA12. These results suggest that it could be useful to break down sectoral European data in regions corresponding to the different member countries. Beck, Hubrich, and Marcellino (2011) propose this approach in the context of price setting.

In any case, the implementation of the double-disaggregation criterion is aimed to make use of the models and forecasts at the level of sectors within a geographical area. In order to show that these highly disaggregated forecasts are trustworthy in both cases – that of Spain and the EA12 – we demonstrate that the accuracy of the resulting headline inflation indirect forecast is not significantly worse than that of other forecasts based on simpler breakdowns analysed in the article. This is relevant because it points out that expanding the analysis from 57/50 aggregated-sector series to 969/600 sectors through all the geographical areas, we do not get worse aggregated results, and we can provide much broader forecasting information through sectors within areas. This is of special interest when putting the results for the highly disaggregated series in terms relative to the global economy under study or to other geographical areas. Espasa and Albacete (2007) and Espasa and Mayo-Burgos (2013) show that when analysing disaggregated data, it is important to consider restrictions between the disaggregates coming from the presence of common features between them. In this article, we study the relevance of spatial cointegration. On the question of how to define neighbourhood, the article provides evidence that the different definitions do not lead to significantly different forecasting results and that considering the whole area under study as the 'neighbourhood' is as good as any definition based on specific geographical areas. This can considerably simplify the treatment of spatial cointegration in these contexts.

Including spatial cointegration restrictions does not significantly improve the aggregate forecast, but it plays a useful role in forecasting inflation in the aggregated sectors, as it can allow us to identify the sectors in which inflationary pressures are more likely to occur.

Regarding modelling and forecasting inflation at the sectoral regional level, we deal here with many relevant questions and show that the results are reliable and, therefore, can be useful by policy makers, investors and agencies watching competitiveness. Other applications seem interesting, as well – for example, to find out how this approach can be used in big countries like the US, where the economic integration of the states could be something in between that of the EA12 member countries and the regions in Spain. Another question for future study involves cointegration through sectors, as in Espasa and Mayo-Burgos (2013), and spatial cointegration together, as well as the consideration of other features besides common trends. Finally, the application to other economic indicators, such as industrial production, also seems promising.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Antoni Espasa and J.D. Tena acknowledge financial support from Ministerio de Educación y Ciencia [grant numbers ECO2009-08100 and ECO2012-32401].

References

Aron, J., and J. Muellbauer. 2012. "Improving Forecasting in an Emerging Economy, South Africa: Changing Trends, Long Run Restrictions and Disaggregation." *International Journal of Forecasting* 28 (2): 456–476. doi:10.1016/j.ijforecast.2011.05.004.

Beaulieu, J. J., and J. A. Miron. 1993. "Seasonal Unit Roots in Aggregate U.S Data." *Journal of Econometrics* 55 (1–2): 305–328. doi:10.1016/0304-4076(93)90018-Z.

Beck, G., K. Hubrich, and M. Marcellino. 2009. "Regional Inflation Dynamics within and across Euro Area Countries and a Comparison with the United States." *Economic Policy* 24 (57): 141–184. doi:10.1111/ecop.2009.24.issue-57.

Beck, G., K. Hubrich, and M. Marcellino. 2011. *On the Importance of Sectoral and Regional Shocks for Price-Setting*. CEPR Discussion Papers 8357, C. E.P.R. Discussion Papers. Washington, DC: Centre for Economic Policy Research.

Boivin, J., and S. Ng. 2006. "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics* 132 (1): 169–194. doi:10.1016/j.jeconom.2005.01.027.

Christoffersen, P. F., and F. X. Diebold. 1998. "Cointegration and Long-Horizon Forecasting." *Journal of Business and Economic Statistics* 16 (4): 450–458.

Diebold, F., and R. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics* 13: 253–265. Published again in *Journal of Business and Economic Statistics* 20 (1): 135–145, January 2002.

Espasa, A., and R. Albacete. 2007. "Econometric Modelling for Short-Term Inflation Forecasting in the Euro Area." *Journal of Forecasting* 26: 303–316. doi:10.1002/(ISSN)1099-131X.

Espasa, A., and I. Mayo-Burgos. 2013. "Forecasting Aggregates and Disaggregates with Common Features." *International Journal of Forecasting* 29 (4): 718–732. doi:10.1016/j.ijforecast.2012.10.004.

Faust, J., J. Rogers, and J. Wright. 2005. "News and Noise in G-7 GDP Announcements." *Journal of Money, Credit and Banking* 37 (3): 403–419. doi:10.1353/mcb.2005.0029.

Franses, P. H. 1991. "Seasonality, Nonstationarity and the Forecasting of Monthly Time Series." *International Journal of Forecasting* 7 (2): 199–208. doi:10.1016/0169-2070(91)90054-Y.

Giacomini, R., and C. W. J. Granger. 2004. "Aggregation of Space-Time Processes." *Journal of Econometrics* 118 (1–2): 7–26. doi:10.1016/S0304-4076(03)00132-5.

Gomez, V., and A. Maravall. 1996. *Programs TRAMO and SEATS, Instruction for Users (Beta version: September 1996)*. Madrid: Banco de España. WP9628.

Harvey, D., S. Leybourne, and P. Newbold. 1997. "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting* 13 (2): 281–291. doi:10.1016/S0169-2070(96)00719-4.

Hasza, D. P., and W. A. Fuller. 1982. "Testing for Nonstationary Parameter Specifications in Seasonal Time Series Models." *The Annals of Statistics* 10 (4): 1209–1216. doi:10.1214/aos/1176345985.

Hendry, D. F. 2006. "Robustifying Forecasts from Equilibrium-Correction Systems." *Journal of Econometrics* 135 (1–2): 399–426. doi:10.1016/j.jeconom.2005.07.029.

Hendry, D. F., and K. Hubrich. 2011. "Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate." *Journal of Business & Economic Statistics* 29 (2): 216–227. doi:10.1198/jbes.2009.07112.

Imbs, J., H. Mumtaz, M. O. Ravn, and H. Rey. 2005. "PPP Strikes Back: Aggregation and the Real Exchange Rate." *Quarterly Journal of Economics* 120 (1): 1–44.

Johansen, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Osborn, D. R., A. P. L. Chui, J. P. Smith, and C. R. Birchenhall. 1988. "Seasonality and the Order of Integration for Consumption." *Oxford Bulletin of Economics and Statistics* 50 (4): 361–377. doi:10.1111/j.1468-0084.1988.mp50004002.x.

Pesaran, M. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74 (4): 967–1012. doi:10.1111/ecta.2006.74.issue-4.

Pesaran, M. H., and R. Smith. 1995. "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels." *Journal of Econometrics* 68 (1): 79–113. doi:10.1016/0304-4076(94)01644-F.

Schumacher, C., and J. Breitung. 2008. "Real-Time Forecasting of German GDP Based on a Large Factor Model with Monthly and Quarterly Data." *International Journal of Forecasting* 24 (3): 386–398. doi:10.1016/j.ijforecast.2008.03.008.

Tena, J. D., A. Espasa, and G. Pino. 2010. "Forecasting Spanish Inflation Using the Maximum Disaggregation Level by Sectors and Geographical Areas." *International Regional Science Review* 33 (2): 181–204. doi:10.1177/0160017609336629.

West, K. D. 2006. "Forecast Evaluation." In *Handbook of Economic Forecasting*, edited by G. Elliott, C. W. J. Granger, and A. Timmermann. Amsterdam: Elsevier.

Appendix. Time series

We use time series for the following disaggregate products in the case of Spain: S1: Cereals; S2: Bread; S3: Beef; S4: Lamb; S5: Pork; S6: Bird; S7: Other meat; S8: Fish; S9: Crustaceans, molluscs and processed fish; S10: Eggs; S11: Milk; S12: Milk products; S13: Oil and fats; S14: Fresh fruit; S15: Preserved fruit; S16: Vegetables; S17: Preserved vegetables; S18: Potatoes; S19: Coffee, cacao and infusions; S20: Sugar; S21: Other food products; S22: Non-alcoholic drinks; S23: Alcoholic drinks; S24: Tobacco; S25: Men's clothes; S26: Women's clothes; S27: Clothes for babies and children; S28: Complements and Repairs; S29: Men's footwear; S30: Women's footwear; S31: Footwear for babies and children; S32: Repair of footwear; S33: Rented apartments; S34: Heating, lighting and water distribution; S35: Own apartments; S36: Furniture and floor coverings; S37: Textile and home accessories; S38: Major appliances; S39: Household items; S40: Non durable household items; S41: Home services; S42: Medical services; S43: Medicines and other chemical products; S44: Personal transportation; S45: Public urban transportation; S46: Public intercity transportation; S47: Mail and communications; S48: Recreational objects; S49: Publications; S50: Recreation; S51: Primary school; S52: Secondary school; S53: University; S54: Other expenditures in education; S55: Personal items; S56: Tourism and hotels; and S57: Other goods and services.

We use time series for the following disaggregate products in the case of EA12: S1: Bread and cereals; S2: Meat; S3: Fish and seafood; S4: Milk, cheese and eggs; S5: Oils and fats; S6: Fruit; S7: Vegetables; S8: Sugar, jam, honey, chocolate and confectionery; S9: Food products n.e.c; S10: Coffee, tea and cocoa; S11: Mineral waters, soft drinks, fruit and vegetable juices; S12: Spirits; S13: Wine; S14: Beer; S15: Tobacco; S16: Clothing; S17: Footwear including repair; S18: Actual rentals for housing; S19: Maintenance and repair of the dwelling; S20: Water supply and miscellaneous services relating to the dwelling; S21: Electricity, gas and other fuels; S22: Furniture and furnishings, carpets and other floor coverings; S23: Household textiles; S24: Major household appliances whether electric or not and small electric household appliances; S25: Repair of household appliances; S26: Glassware, tableware and household utensils; S27: Tools and equipment for house and garden; S28: Non-durable household goods; S29: Domestic services and household services; S30: Health; S31: Motor cars; S32: Motor cycles, bicycles and animal drawn vehicles; S33: Spares parts and accessories for personal transport equipment; S34: Fuels and lubricants for personal transport equipment; S35: Maintenance and repair of personal transport equipment; S36: Other services in respect of personal transport equipment; S37: Transport services; S38: Postal services; S39: Telephone and telefax equipment and services; S40: Audio-visual, photographic and information processing equipment; S41: Other major durables for recreation and culture; S42: Other recreational items and equipment, gardens and pets; S43: Recreational and cultural services; S44: Newspapers, books and stationery; S45: Package holidays; S46: Education; S47: Restaurants, cafés and the like; S48: Canteens; S49: Accommodation services; S50: Miscellaneous goods and services.